Team Reports: Marine Biological Laboratory

Marine Biological Laboratory
Executive Summary
Principal Investigator: Mitchell Sogin

From Early Biospheric Metabolisms to the Evolution of Complex Systems

Early life forms:

One central theme in Astrobiology is to investigate how ancient life forms influenced early Earth with reference to the formation of habitats capable of supporting complex biological communities and multi–cellular organisms. Based upon what we know about Earth's 2.2–3.5–billion–year evolutionary history, microbial organisms were its sole inhabitants until the origins of plants and animals, a mere 560—900 million years ago. The general goal of the Astrobiology Program at the Marine Biological Laboratory (MBL) is to investigate early evolving metabolisms and activities that had the potential to reshape planetary environments, and to understand the evolution of genome architecture that led to more complicated life forms. These studies contribute to flight–related missions through the development of life detection technologies and design of models for remote sensing of bio–signatures. Many of our projects focus on the evolution of single–cell organisms and complex microbial communities that live under extreme conditions including anoxic and iron–rich environments. Others use microbial model systems (both free–living and symbionts of metazoans) to understand the evolution of metabolic innovation and mechanisms that govern genome evolution. Our genomic, microbial and metabolic diversity–related investigations take advantage of high–throughput DNA sequencing capacity at the MBL and powerful computational systems for phylogenetic inference.

The importance of microbial–centric studies of astrobiology reflects the key role that these organisms have always played in biogeochemical processes on Earth. Microbes of untold diversity are the primary catalysts of energy transformation, and are responsible for > 98% of the carbon and nitrogen cycling [1]. An estimated 3.6×10^{30} microbial cells with cellular carbon of ~3×10^{17} grams may account for more than 90 percent of the total biomass [2]. Given the ability of microbes to transform Earth's environment from one that was highly reducing to a habitat that is capable of supporting complex, oxygen breathing, multi–cellular organisms, we hypothesize that life on other planets will also be microbial in form. The general goal is to understand patterns and mechanisms of genome evolution and metabolic variation that allowed diverse

microorganisms to adapt to new environments, generate novel phenotypes, and evolve processes that led to environmental changes on a global scale, some of which can be detected through remote sensing.

Metabolisms in extreme environments:

Our investigations of early metabolic evolution focus upon phylogenetic analysis of functional genes from hydrothermally altered anoxic sediments of Guaymas basin in the Gulf of California, the Deep subsurface from the Peru Margin, the acidic, iron–rich, heavy metal laden Río Tinto of southwestern Spain, and cultured isolates from anoxic marine sediments or the Juan de Fuca ridge.

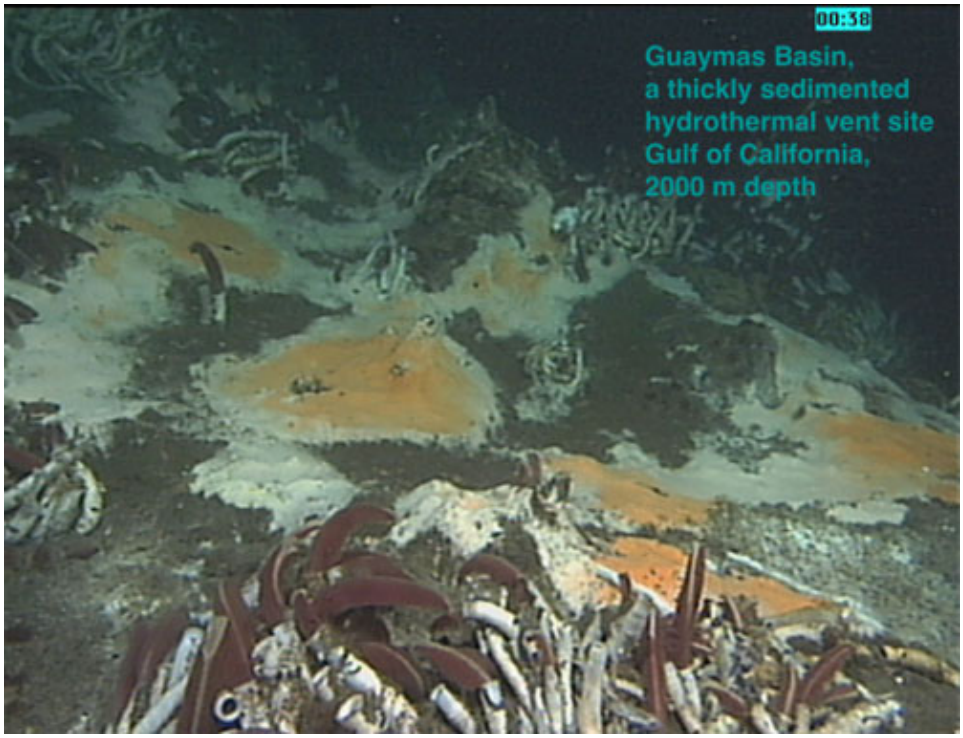

Figure 1. Guaymas Basin Ecosystem. Hydrothermally active sediments in Guaymas Basin and Yellow and Orange Begginton mats grow where sulfidic vent fluid diffuses up from the sediment. Anaerobic CH_4 and S cycling in these sediments is catalyzed by a diverse community of sulfate reducers, methanogen archaea and methanotrophic consortia, including members of novel, deeply branching lineages.

These environments most likely reflect conditions on early Earth or a wetter Mars. Life on early Earth developed within an anaerobic environment and evolved physiologies and metabolisms that were ancestral to modern–day sulfate–reducing, methanogenic, and methane–oxidizing microbes. Our initial surveys of functional genes targeted the anaerobic pathways of sulfate reduction (dissimilatory sulfite reductase (*dsr*)) [3] and methanogenesis (coenzyme M methyl reductase (*cmr*)). We are expanding these investigations to include genes associated with the reductive TCA cycle and iron oxidation. All of these investigations rely upon microbial population structure analyses

(based upon sequence analysis of polymerase chain reaction products for rRNA genes) to link evolutionary studies of functional genes with specific classes of microorganisms (Palacios, Amaral Zettler and Sogin, manuscript in preparation).

Microbial population structures in extreme environments: Based upon full length rRNA sequence analyses we previously reported unexpected eukaryl diversity in both the Guaymas basin and Río Tinto [4, 5].

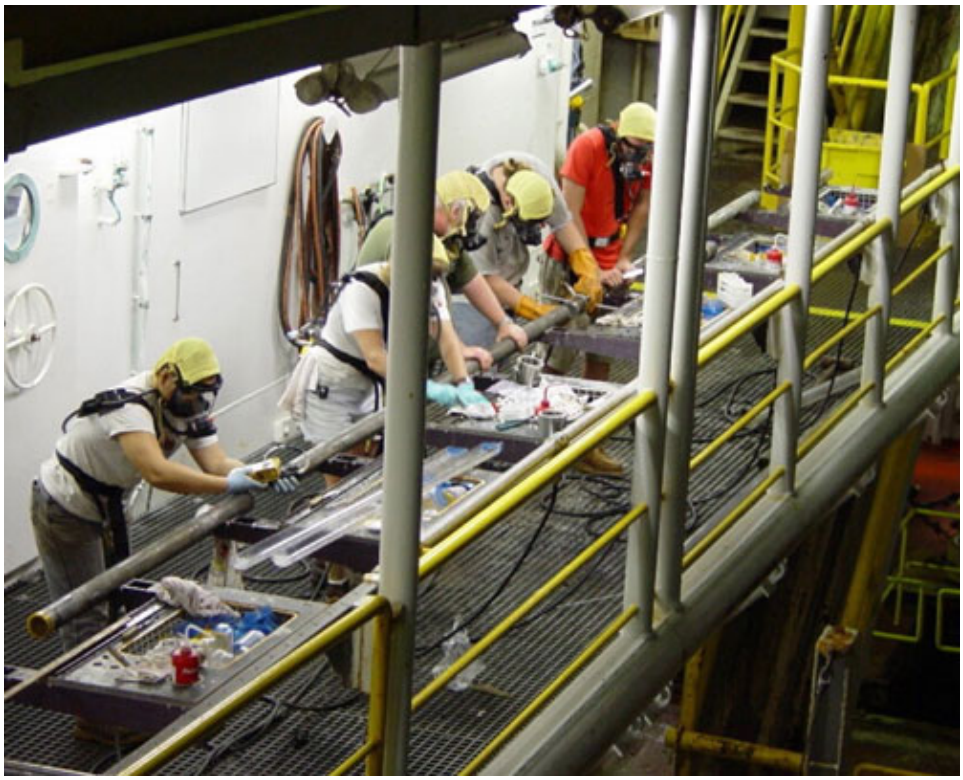

Figure 2. Night sectioning of a freshly retrieved deep sediment core from the Peru Margin. The core is rich in the poisonous gas H2S that is produced by sulfate–reducing bacteria that are active in the deep subsurface; protective masks have to be worn by the shipboard technicians. The subsurface microbiota in these organic–rich cores (sites 1227 to 1230) were investigated with multidisciplinary approaches (molecular biology, dominated by subsurface lineages (Chloroflexi phylum, JS1 phylum; DSAG archaeal lineage) that occur consistently in the subsurface.

In contrast, bacterial diversity is much lower with polymerase chain reaction (PCR) amplicons representing only about 1/3 of the known phyletic diversity of bacteria in Guaymas and only a handful of different kinds of bacteria from the Río Tinto. To expand these descriptions of microbial population structure Kysela, Palacios, and Sogin developed a high throughput technology, serial analysis of ribosomal sequence tags of the V6 (SARST–V6) [6], for characterizing microbial population structures at the DNA sequence level. With this technology we capture sequence information from many more microbial taxa at 1/10th the cost of more traditional sequence analyses of PCR amplicons for full–length rRNA genes.

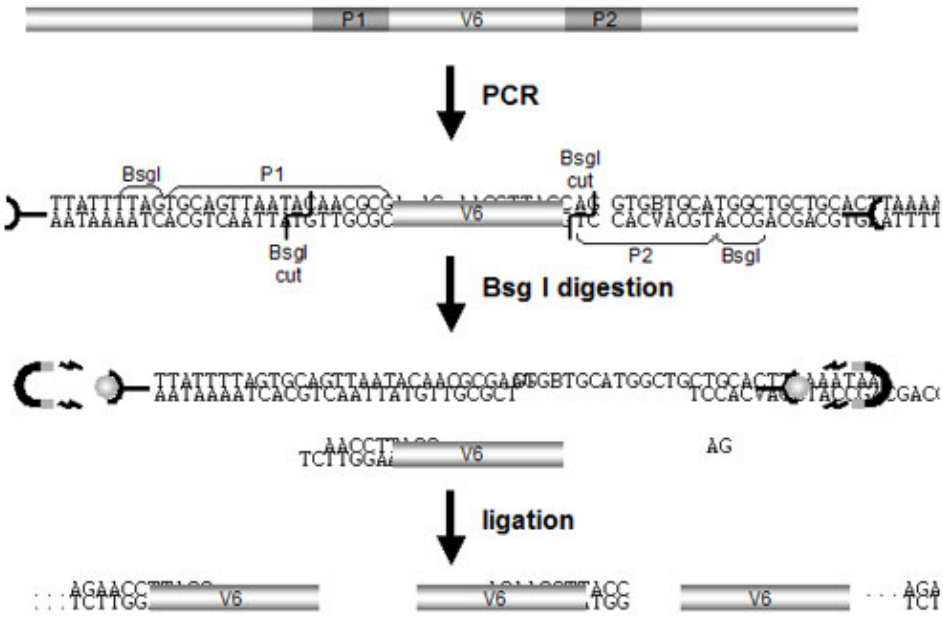

Figure 3. Overview of the serial analysis of V6 ribosomal sequence tags (SARST–V6) method. Biotinylated PCR primers target conserved sites flanking the V6 hypervariable region of the 16S rRNA gene. Primers include a 5 linker sequence containing the recognition sequence for the type IIs restriction enzyme BsgI. Digestion by BsgI results in DNA cleavage 16 bp downstream of the recognition sequence on the sense strand and 14 bp downstream on the antisense strand, leaving a 2 bp overhang at each end of the amplicon. Digested termini are purified away using magnetic streptavidin–coated beads. Ligation of digested amplicons yields concatemers with multiple, serially arranged PCR products. Sequence regions are not to scale.

In brief, SARST–V6 produces sequences of large concatemers of PCR–amplified ribosomal sequence tags (RSTs) from homologous V6 hypervariable regions in rRNA genes. Comparison against a comprehensive rRNA gene database identifies the taxonomic assignment of individual RSTs in the concatemers. SARST–V6 thus evaluates the diversity and relative numbers of different rDNA amplicons from a heterogeneous microbial population. Since SARST–V6 allows for the sampling of 10–20 rRNA genes by sequencing both ends of a single cloned PCR amplicon, we obtain improved statistics for descriptions of different kinds of microbial taxa in environmental DNA samples. There was excellent correspondence between the SARST–V6 analysis and full–length rRNA studies from Guaymas basin cores. In the case of the Río Tinto, SARST–V6 reported all of the microbial taxa described using more traditional microbial ecological methods. It also identified new microbes that were not detected in full–length sequence analyses of PCR amplicons of rRNAs from the Río Tinto including taxa observed in geographically remote acidic environments such as Iron Mountain.

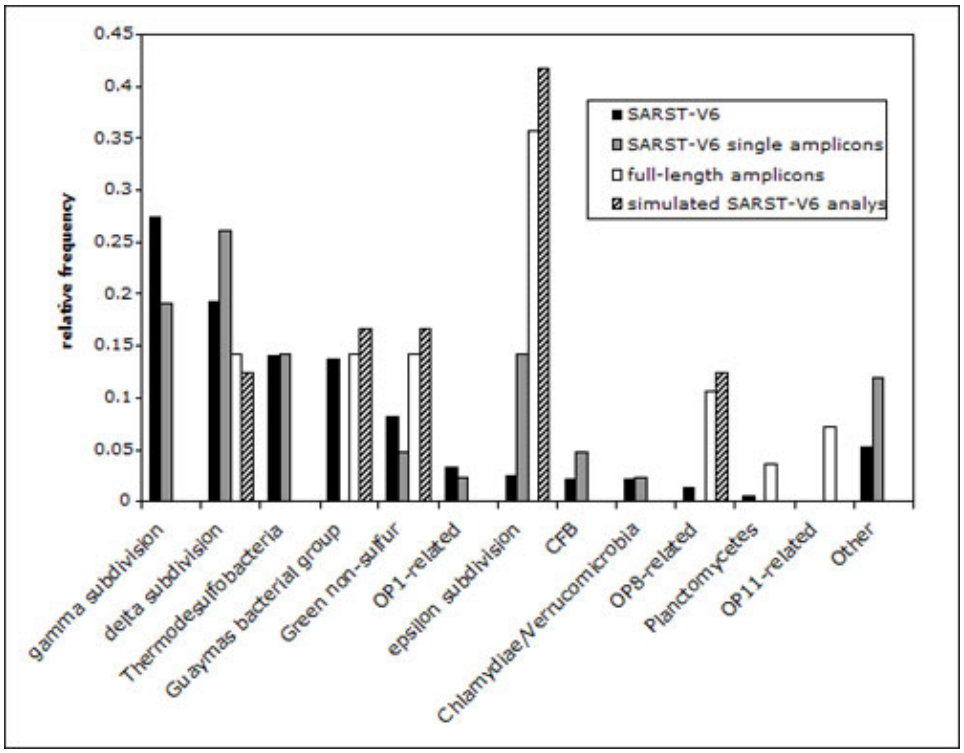

Figure 4. Bacterial phylotype distributions obtained using SARST−V6 and other analysis methods. All data were collected on hydrothermally heated sediments from the Guaymas Basin (Gulf of California). Results of four analyses are shown: SARST−V6 (black; N=364), single amplicon cloning using SARST−V6 primers (gray, N=42), phylogenetic analysis of full−length 16S rRNA gene clone sequences (white, N=28) and a simulated SARST−V6 analysis of full−length sequences (hatched, N=24). Except for phylogenetic analysis of full−length gene sequences, taxonomic assignments were based on results of BLAST searches of the Genbank nt database.

Sulfite Reducatases: We have previously reported on the diversity of the dissimilatory sulfate reduction genes (*dsrAB*) [3] and recently included those sequences in a phylogenetic analysis of conserved active centers from both dissimilatory and assimilatory sulfite reductases [7]. Dhillon and Goswami demonstrated that these activities are phylogenetically linked through an ancestral monomeric reductase. A gene duplication of the reductase generated the assimilatory and dissimilarity forms with the latter undergoing a second gene duplication for form dsrA and dsrB prior to the split between bacteria and archaea. We have now expanded the surveys of functional genes through studies of mcrA, the alpha subunit of methyl−Coenzyme M reductase.

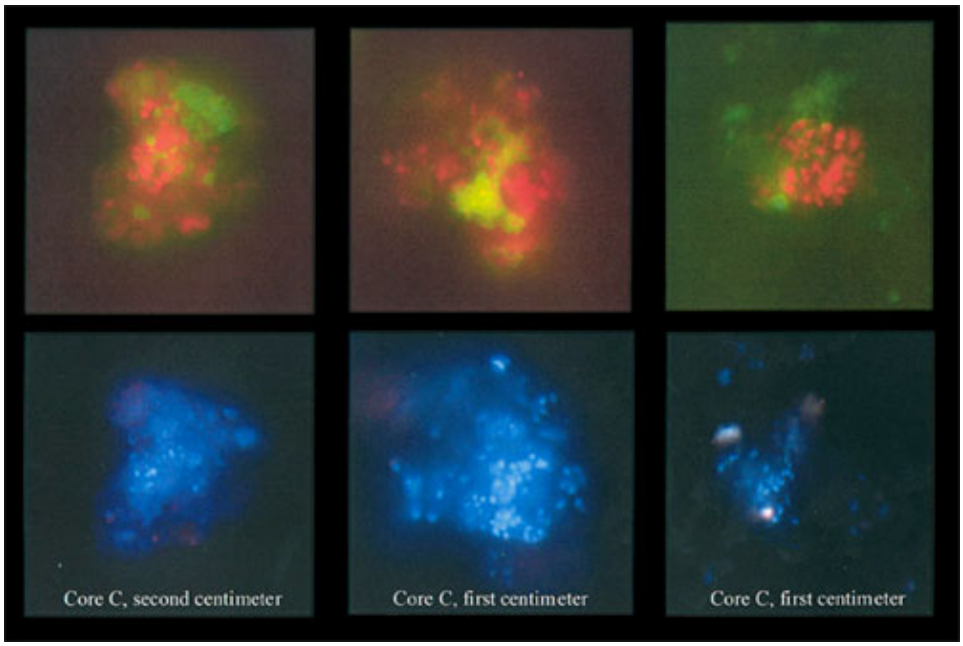

Figure 5. FISH (Fluorescent in–situ hybridization) with 16S rRNA–targeted oligo nucleotides of archael/bacterial consortia in Guaymas Basin. Archaea are stained red, bacteria green, DAPI stained images are blue. These consortia are the likely catalysts of Anaerobic methane oxidation in Guaymas. Photo courtesy of K. Knittel and A. Boetius.

We have identified within environmental DNA samples from Guaymas sediments mcrA genes from the methanogens Methanosarcinales, Methanomicrobiales, and Methanococcales. There appears to be a diversified, partially thermophilic methanogen community in guaymas with several uncultured lineages among the Methanosarcinales and Methanomicrobiales. Within the next year, we anticipate the development of a more comprehensive description of phylogeny and divergence of methanogenic terminal anaerobic degraders based upon evolutionary analyses of several functional genes and rRNA coding regions.

Reductive TCA metabolism: Autotrophic carbon fixation under anaerobic conditions represents yet another putatively ancient pathway for early evolving bacteria. The reductive TCA cycle is essentially the TCA cycle running in reverse, leading to the fixation of 3 molecules of CO_2 and the production of one molecule of triose phosphate. Most of the enzymes between the two pathways are shared, with the exception of two key enzymes that allow the cycle to run in reverse: ATP citrate lyase (ACL) and 2–oxoglutarate:ferredoxin oxidoreductase (OOR). To explore the phylogeny of ACL and OOR, Sievert and Edgcomb selected thirteen thermophillic bacterial and archaeal strains suspected of carrying out reductive for physiological and molecular analyses. Using ten different kinds of media under multiple anaerobic incubation temperatures, they obtained biochemical evidence for the operation of the reductive TCA cycle in two of the e–proteobacteria, and in *Desulfurobacterium autotrophicum* (*Aquificales*). Custom designed primers allowed for the PCR amplification of ACL genes from several cultures and from this limited data set there is congruence between trees inferred from 16S and ACL genes. This suggests that ACL will be a valuable tool for exploring RTCA evolution from

environmental DNA samples isolated from different chemosynthetic environments.

Iron metabolism and iron rich environments: Iron has always played an important role in life's history. Banded iron formations (BIF) date back to the Archean, and proteins with FeS centers are widely distributed in the tree of life. Several of our projects share a common interest in microbial Fe metabolism or growth in Fe−rich environments such as the Río Tinto. The study of environments enriched in iron may provide insights into the biochemistry and metabolisms of the early Earth and ancient Mars. Despite the pivotal role of iron in biological systems, there is only limited knowledge about Fe oxidation as an ancient metabolic pathway. Edwards has isolated several new Fe oxidizing autotrophs from deep−sea cultures.

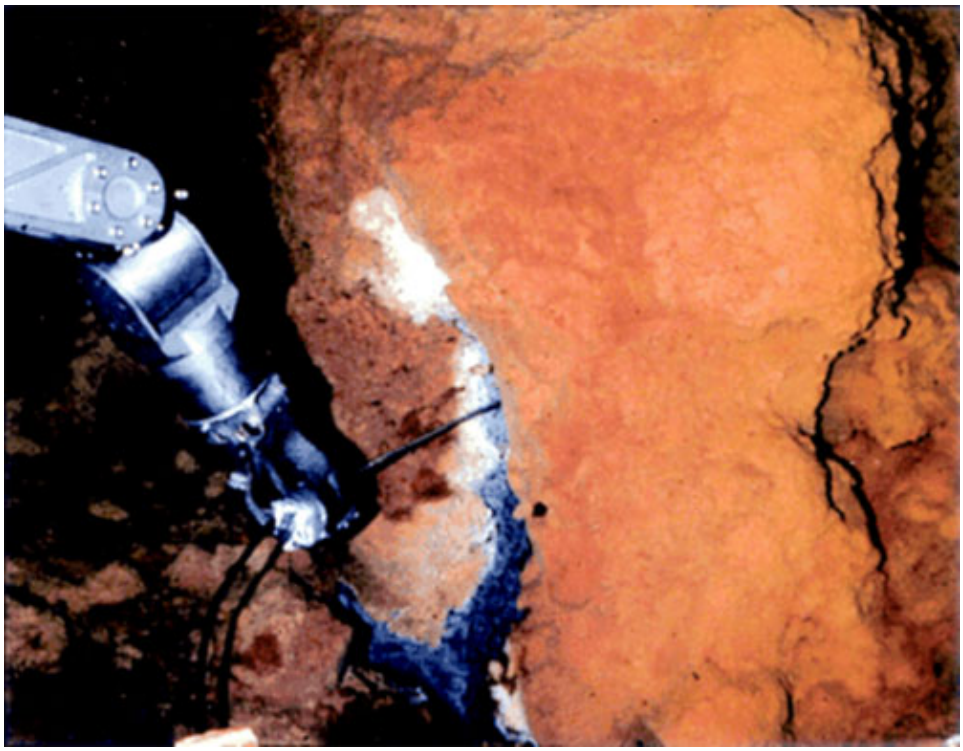

Figure 6. Loihi hydrothermal vent site. A temperature probe is being inserted into the vent orifice; note the extensive deposits of Fe oxides surrounding the vent opening.

These strains will provide an opportunity to explore phylogenetic diversity of neutrophillic Fe metabolism and to establish model systems for molecular metabolic studies. The isolates are facultative autotrophs and/or mixotrophs and most are phylogenetically affiliated with either gamma or epsilon proteobacteria. *Marinobacter aquaeoloi* serves as a model system for exploring Fe metabolism. It has the growth characteristics of a psychrophilic bacterium with a 35 kDa protein that is up regulated with increasing Fe(II) content in medium. Dhillon, Sogin and Edwards have constructed a shot−gun genomic library that will be used for high throughput DNA sequencing. We have submitted a request to the Department of Energy's Joint Genome Institute (JGI) for full genome sequencing of the library. However, if the JGI is unable to sequence the M. aquaeoloi genome within the next few months, we will take

advantage of the low−cost, high−throughput facilities in Woods Hole to ensure the rapid analysis of the *M. aquaeoloi* genome. To identify genes involved in iron metabolism, Edwards has developed a novel plate−based screening assay that relies upon formation of colored colonies for the detection of Fe oxidation. This assay will facilitate the optimization of conditions for transposon mutagenesis and for the cloning of genes involved in Fe oxidation into *E. coli*. Using a 40−kb DNA fosmid insert library from *M. aquaeoloi*, Dhillon has transformed *E. coli* and the observed changes in colony coloration (pinkish−brown colonies) are consistent with development of functional iron oxidation under neutral conditions in the presence of ferric chloride. Dillon and Edwards will use the same assay to identify recombinant clones with iron oxidation activities from large insert libraries of environmental DNAs extracted from Fe−oxidizing mats. In a similar manner, Amaral Zettler and Edwards will isolate eukaryotic genes that play a role in iron oxidiation from the Río Tinto.

The development of this plate−based functional assay for iron metabolism will play an important role in a new project to be carried out by Julie Huber who recently received an NRC/NAI Post Doctoral award. Her project will explore the genomic context of different functional genes in geographically distinct environments that have similar, if not identical, water chemistries. While molecular and culture studies have revealed a phylogenetically and physiologically diverse microbial community indigenous to the subseafloor, very little is known about either its genetic content or its functional and metabolic potential. Genome studies suggest that functional genes can shuttle between microbial lineages. If early life experienced higher levels of lateral gene transfer than contemporary communities, this may have allowed microbes to reassemble or borrow genes, thus facilitating adaptation into diverse anaerobic thermal habitats. Whether similar processes are dominant factors in reshaping adaptation by microbial populations today is unknown. With the development of functional assays that can interrogate large insert libraries prepared from environmental DNA, we predict that it will be possible to define the role of horizontal gene transfer in natural settings. The ability to identify genes responsible for Fe oxidation according to function rather than sequence similarity (as is commonly done for PCR−based studies) will allow us to examine the genomic context of genes that have a clearly defined function.

A Fe−rich Terrestrial Analogue for Early Mars – Río Tinto: Investigations of microbial diversity and population structure in the Río Tinto will link shifts in microbial communities with changes in

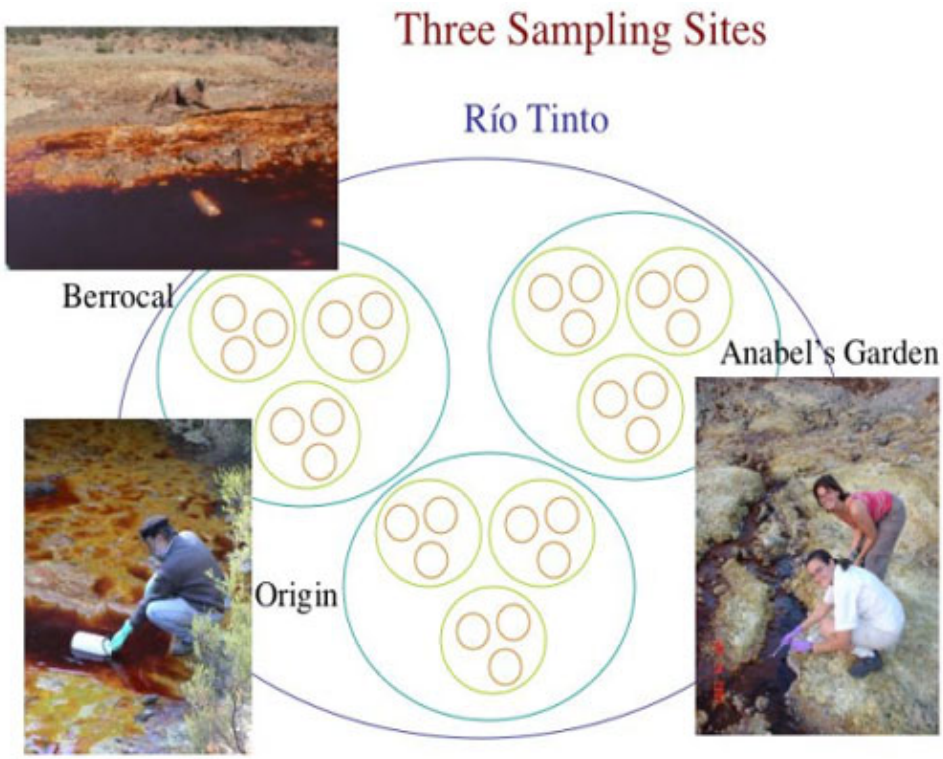

Figure 7. Sampling strategies for microbial population structure studies in the Rio Tinto. The Origin is generally more acidic and has higher iron concentrations (>20 mg/ml) than other locations. Anabel's Garden supports large biofilm communities and elevated levels of ferrous iron. Berrocal has the highest flow rates and is sometime anoxic.

water chemistry. Two sampling trips, the fall of 2003 and January of 2004, correspond to the dry and wet seasons. Using three stations where different oxidiation states of iron are present, Amaral Zettler and Palacios of the MBL and Ricardo Amils of Centro de Astrobiologia (CAB) have sampled DNA, collected physico–chemical parameters consisting of *in situ* (pH, oxygen, redox, temperature and conductivity) and *ex situ* (S, Fe, Zn, Cu, Al, As, Ni, Mg, Ca, K, etc.) measurements, and preserved formalin samples for bacterial and protist enumeration. Using SARST–V6 described above, we are developing in–depth studies of microbial population structure for both eukaryotic and prokaryotic organisms. In addition to collection of 70 full–length rRNA PCR amplicon sequences, Palacios has also recovered more than 4,000 SARST sequenced tags for calculating alpha and beta diversity indices. The SARST diversity data will be combined with the physico–chemical parameters in analyses that will reveal important aspects of the biogeochemistry of this extreme environment. To date, the analysis of the 70 full–length sequences revealed the presence of phylotypes not detected in previous studies including members of the archaeal “alphabet” plasmas, close bacterial relatives to taxa found in other acidic extreme environments and examples of anaerobic amoeboid protists. The SARST–V6 analysis

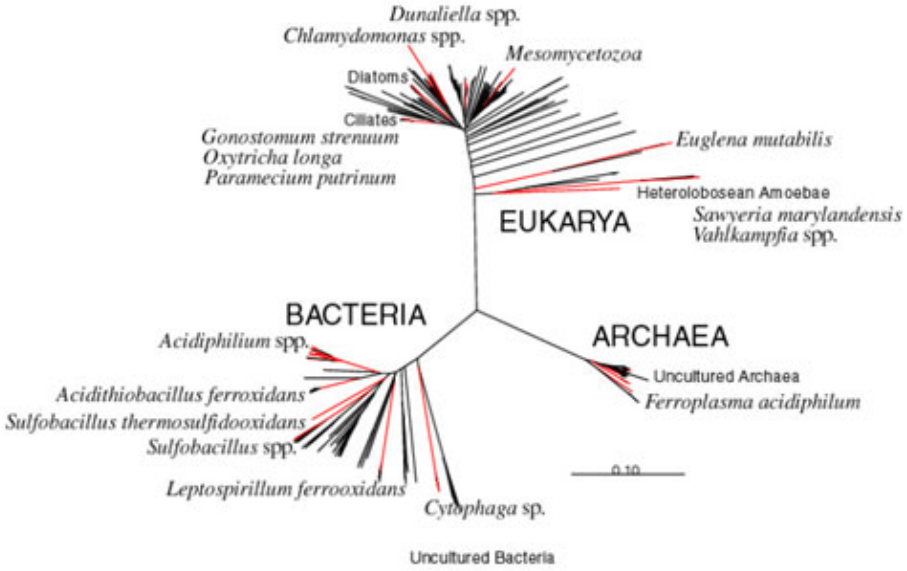

Figure 8. Phylogenetic tree of nearly full–length rRNA genes isolated from the Rio Tinto sampling sites. The red lineages indicate the phylogenetic placement of a limited number of small subunit rRNA environmental clones derived from specific bacterial, archaeal and eukaryotic PCR amplicon clone libraries. The actual phylogenetic affinity of the environmental clones is indicated by labeled lineages. This tree shows the minimal range of phylogenetic diversity of the three domains of life present in the Rio Tinto.

detected the same bacteria in the Río Tinto as reported using other microbial ecology methods, plus candidate novel organisms for extreme environments including free–living bacteria, parasites and endosymbionts. In conjunction with the microbial population structure studies and biogeochemical measurements for the Río Tinto sampling sites, the geologists in our NAI team are developing a strategy for remote sensing of the Río Tinto that will be applicable to Mars. The development of spectral libraries and databases required for interpreting remote sensing data will be completed next year and will be correlated with shifts in microbial populations and biogeochemicals. The goal is to use the Rio Tinto as a development site for interpreting remote sensing data from the Mars Express instruments.

Microbial symbionts and the evolution of genome architecture:
Evolutionary comparisons of genome sequences (phylogenomics) provide compelling evidence that horizontal gene transfer serves a major role in the evolution of genome architecture. Symbioses in which an organism inhabits the cytoplasm of a second organisms has the potential to accelerate horizontal gene flow and to influence genome evolution. The origins of mitochondria and chloroplasts are well–documented examples. These organelles were derived from alpha proteobacterial (mitochondria) or cyanobacterial (plastids) symbionts. The genomes of these organelles experienced major change including shifts in base composition, reductions in genome complexity, and transfer of key coding regions into the host genome. These events occurred over long evolutionary time frames, but we know little about underlying mechanisms. In contrast, studies of contemporary bacterial symbioses in animal hosts offer an opportunity to explore mechanisms that might explain the

remodeling of symbiont and host genome architecture. Wernegreen's and Bordenstein's genome studies of divergent bacterial symbionts in insects have shown there is retention of genes according to the nutritional physiology of specific hosts and they offer hypotheses to explain radical shifts in base composition for endosymbiont genomes [8]. The movement of genes also involves newly discovered bacteriophages of Wolbachia which appear to move laterally between divergent strains, and may shuttle chromosomal genes [9]. Sequence analysis is now complete for the full genome of *Blochmannia* associated with *Camponotus pennsylvanicus* (i.e., *B. pennsylvanicus)*, totaling 19,480 reads (15.4 Mb total) from a short–insert genomic library to obtain 12–fold coverage of this 792 kb chromosome and efforts to annotate this closed genome are underway with Dr. Monica Riley's group.

1. Atlas, R.M. and Bartha, R. *Microbial Ecology. Fundamentals and Applications*. 1993, Redwood City: The Benjamin/Cummings Publishing Company, Inc.

2. Whitman, W.B., Coleman, D.C., and Wiebe, W.J. *Prokaryotes: The Unseen Majority.* Proc Natl Acad Sci, U.S.A., 1998. **95**(12): p. 6578–83.

3. Dhillon, A., Teske, A., Dillon, J., Stahl, D.A., and Sogin, M.L. *MolecularCcharacterization of Sulfate–reducing Bacteria in the Guaymas Basin.* Appl Environ Microbiol, 2003. **69**(5): p. 2765–72.

4. Teske, A., Hinrichs, K.U., Edgcomb, V., de Vera Gomez, A., Kysela, D., Sylva, S.P., Sogin, M.L., and Jannasch, H.W. *Microbial Diversity of Hydrothermal Sediments in the Guaymas Basin: Evidence for Anaerobic Methanotrophic Communities.* Appl Environ Microbiol, 2002. **68**(4): p. 1994–2007.

5. Amaral Zettler, L.A., Gomez, F., Zettler, E.R., Keenan, B.G., Amils, R., and Sogin, M.L. *Eukaryotic Diversity in Spain's River of Fire.* Nature, 2002. **417**: p. 137.

6. Kysela, D.T., Palacios, C., and Sogin, M.L. *Serial Analysis of V6 Ribosomal Sequence Tags (SARST–V6).* Environmental Microbiology, 2004. **in revision**.

7. Dhillon, A., Goswami, S., Teske, A., Sogin, M.L., and Riley, M. *Domain Evolution and Functional Diversificatin of Sulfite Reductases.* Astrobiology, in review, 2004.

8. Wernegreen, J.J. and Funk, D.J. *Mutation Exposed: A Neutral Explanation for Extreme Base Composition of an Endosymbiont Genome.* Journal of Molecular Evolution, in press, 2004.

9. Bordenstein, S.R. and Wernegreen, J.J. *Bacterialphage Flux in Endosymbionts (Wolbachia): Lateral Transfer, Recombination Rates and the 'Intracellular Arena' Hypothesis.* Molecular Biology and Evolution, in press, 2004.